Transputer self-organizing map algorithm for beam background rejection at the BELLE silicon vertex detector

Jens Sören Lange[a],*, Chikara Fukunaga[a], Manobu Tanaka[b], Andrzej Bożek[b]

[a] Tokyo Metropolitan University, Minami-Osawa 1-1, Hachioji, Tokyo 192-03, Japan
[b] KEK High Energy Accelerator Research Organization, Oho 1-1, Tsukuba, Ibaraki 305, Japan

Abstract

A growing self-organizing map using a gravitational algorithm was implemented on a transputer to study possible separation of GEANT simulated beam background events and physics events ($e^+e^- \rightarrow B^\circ \bar{B}^\circ$, $q\bar{q}$, ggg, $2\gamma, \tau^+\tau^-$). A fraction of 75.0% of beam background events can be rejected, 96.9% of physics events pass the classification. The decision time is $\tau \leq 2$ ms, thus the system could be used online as level 3 trigger as well as for offline data filtering purposes. © 1999 Published by Elsevier Science B.V. All rights reserved.

1. Introduction

The BELLE detector [1] is a large scale detector system for precision measurements of charged and neutral particles, to be implemented at the KEK-B facility [2], a high luminosity asymmetric e^+e^- collider. The main experimental goal will be the investigation of CP violation ($C\!\!\!/P\!\!\!/$) in B meson systems. The Silicon Vertex Detector (SVD) is the most inner subdetector of BELLE, consisting of 81.900 silicon strips with a strip width of 84 μm for z strips and 50 μm for $r\varphi$ strips. The B meson vertex resolution will be $\Delta z \geq 57$ μm.

The source of the B mesons is the decay of the $\Upsilon(4S)$ resonance, produced in an e^+e^- collision at $\sqrt{s} = 10.6$ GeV. The final multiparticle state consists of ~ 10–12 charged particles, detected in the SVD, along with 10–20 neutral particles (γ, v), detected e.g. by the BELLE CsI(Tl) calorimeter. Besides B physics other reactions types are of interest, too, and should pass any trigger system for final data acquisition (cf. Section 2.2).

One of the main background sources is given by beam background reactions (cf. Section 2.3), resulting from collisions of the beam e^+ or e^- with residual gas in the accelerator vacuum pipe. The expected trigger rate is $R_{BG} \geq 100$ Hz, well down below the maximum BELLE DAQ readout rate of $R = 500$ Hz [3]. However, there is a necessity to study a possible beam background rejection algorithm because R_{BG} is a linear function of the residual gas concentration and, hence, of the beampipe pressure P. If the design parameter of KEK-B vacuum system $P = 1 \times 10^{-9}$ torr [2] is not achieved, e.g. during accelerator commissioning phase, much

* Corresponding author. Tel.: + 81 426 77 3025; fax: + 81 426 77 3002; e-mail: soeren@tmubsun.center.metro-u.ac.jp.

higher rates R_{BG} must be expected. As will be pointed out in Section 4.1, beam background events have the highest charge multiplicity and, hence, the largest SVD data size of all event types. Limitations in the bandwidth in DAQ data transfer interfaces therefore lead to the necessity of rejecting beam background events, if possible.

The Data Acquisition (DAQ) system of the SVD (cf. Section 3.2) is proposed to consist of T805 transputers (cf. Section 3.1). In this paper we study the possibility to use the transputer system not only for DAQ purposes, but also as event pattern recognition system for rejection of beam background. The basic design parameters of this system are as follows: (a) only usage of SVD information, (b) rejection of a fraction of physics events less than 5% and (c) the decision time $\tau \leq 2$ ms for a possible use of the system as level 3 trigger (corresponding to $R = 500$ Hz).

The SVD DAQ system is principally separated from other BELLE subdetector DAQ systems, leading to the requirement (a). Moreover, beam background mainly consists of charged, low momentum particles ("spent electrons", cf. Section 2.3). Thus, neutral particle information, measured by other BELLE subdetectors, is principally not necessary for classification.

Requirement (c) itself leads to the necessity of using a fast, simple algorithm. Due to this requirement it was decided to use a 2-dim $z\varphi$ topology

analysis rather than a tracking algorithm. Futhermore, it set emphasis on the inner layer ($r = 3.0$ cm, -4.0 cm $\leq z \leq 7.6$ cm, $z = 0$ is beam collision point) of the 3-layered SVD, because this layer shows a higher occupancy for beam background events in the order of $\sim 10\%$. The corresponding polar angle region is $36.8° \leq \vartheta \leq 158.5°$.

Furthermore, a neural network based algorithm was chosen (cf. Section 4), because the iteration loops can be parallelized easily on hardware systems as transputers because of iterating loop structure, i.e. always identical commands are given. A shower analysis based upon tracking would require an algorithm with many different instructions, thus not easy to parallelize.

2. Event types

2.1. Monte Carlo event simulation

As BELLE will start its operation at the end of 1998, Monte Carlo simulated data were used for this study. Event generation for $B°\bar{B}°$,$q\bar{q}$ hadron production, 3 gluon events and $\tau^+\tau^-$ production by annihilation were performed using the event generator QQ [5]. Two-photon events were generated using the event generator TREPS [6]. A beam background generator was especially developed for KEK-B accelerator geometry by Sahu [7].
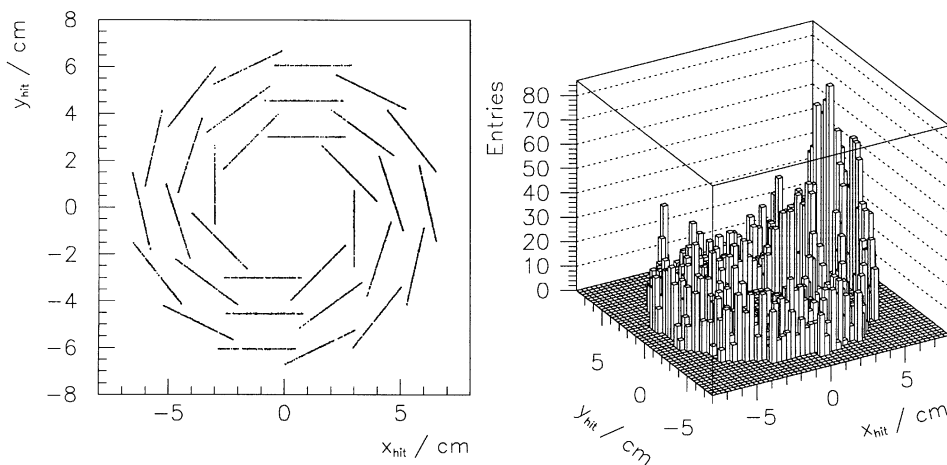

Fig. 1. Hit positions y_{hit} vs. x_{hit} for GEANT simulated beam background events. Left: Beam axis view. The SVD detector structure is visible. Right: Lego plot. An enhancement in the $\varphi \simeq 0$ region is visible.
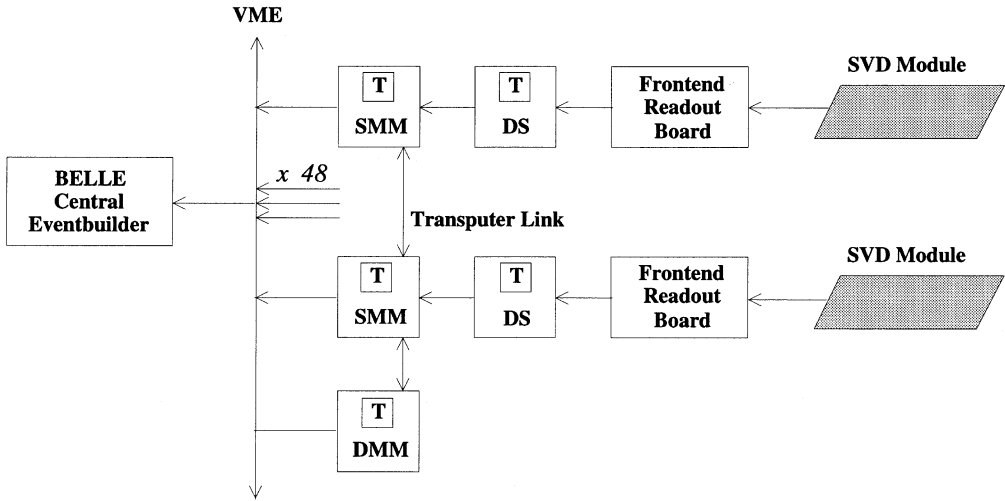
Fig. 2. The proposed SVD DAQ system. Data are transferred from the frontend readout board to the Data Scanner (DS) and further to the Slave Memory Board (SMM). The DMA Master Module (DMM) controls the readout. For details see Section 3.2. Modules containing a T805 transputer chip are indicated with $\boxed{T}$. Totally 103 transputers are planned.

Fig. 1 (left) shows hit positions y_{hit} vs. x_{hit} for simulated beam background events, as an example for all types of simulated events. The 3-layered SVD detector structure becomes visible in this view (beam-eye view along z axis).

Although only charged hits are taken into account for the algorithm, neutral particles are included in the simulation and can generate charged particles by decay or radiation (e.g. $\gamma \rightarrow e^+ e^-$). The SVD detector geometry and response were simulated using the GEANT 3.21 [8] based detector simulation program GSIM, part of the BELLE software library. Data production was performed using the analysis framework BASF [9] with its ability of parallel processing on up to 27 SUN SPARC processors. Beam background simulation typically requires a CPU time of $\tau \simeq 1$ min per event, if simulated on one processor alone. The SVD DAQ system is shown in Fig. 2. It is discussed in detail in Section 3.2.

2.2. Physics events

Five multiparticle reaction types should be considered here to represent all possible $e^+ e^-$ reaction types. They will be referred to as *physics events* hereafter. Generally, an asymmetric choice of beam energies ($T(e^-) = 8$ GeV, $T(e^+) = 3.5$ GeV) results in a non-zero lab system velocity of the $\Upsilon(4S)$ center-of-mass and slightly asymmetric event shapes with respect to the collision point $z = 0$ for all physics event types.

Bhabha scattering $e^+ e^- \rightarrow e^+ e^-$ is not considered in this paper, in spite of the high cross section $\sigma_{Bhabha} = 44$ nb. The two high energetic back-to-back single track signatures will be recognized by the BELLE central driftchamber level 1 trigger with high efficiency.

2.2.1. $B^\circ \overline{B}^\circ$ events

The study of expected ($C\!\!\!/P$) in B meson systems is the main physics task at BELLE. In an $e^+ e^-$ collision at $\sqrt{s} = 10.6$ GeV the $\Upsilon(4S)$ resonance is produced with a cross section of $\sigma_{\Upsilon 4s} = 2.4$ nb. The $\Upsilon(4S)$ decays into $B^\circ \overline{B}^\circ$ and into $B^+ B^-$ with a branching fraction of 50% both. Only $B^\circ \overline{B}^\circ$ events were used in this study, because of the expected higher values[1] of ($C\!\!\!/P$). Decay channels

---

[1] Due to interference of ($C\!\!\!/P$) in decay ($\Gamma(B \rightarrow X) \neq \Gamma(\overline{B} \rightarrow \overline{X})$) and ($C\!\!\!/P$) in oscillation $p(B^\circ \rightarrow \overline{B}^\circ) \neq p(\overline{B}^\circ \rightarrow B^\circ)$ predicted values of ($C\!\!\!/P$) are larger in case of $B^\circ \overline{B}^\circ$ than in case of $B^+ B^-$ (with decay widths Γ, decay channels X and probabilities p).
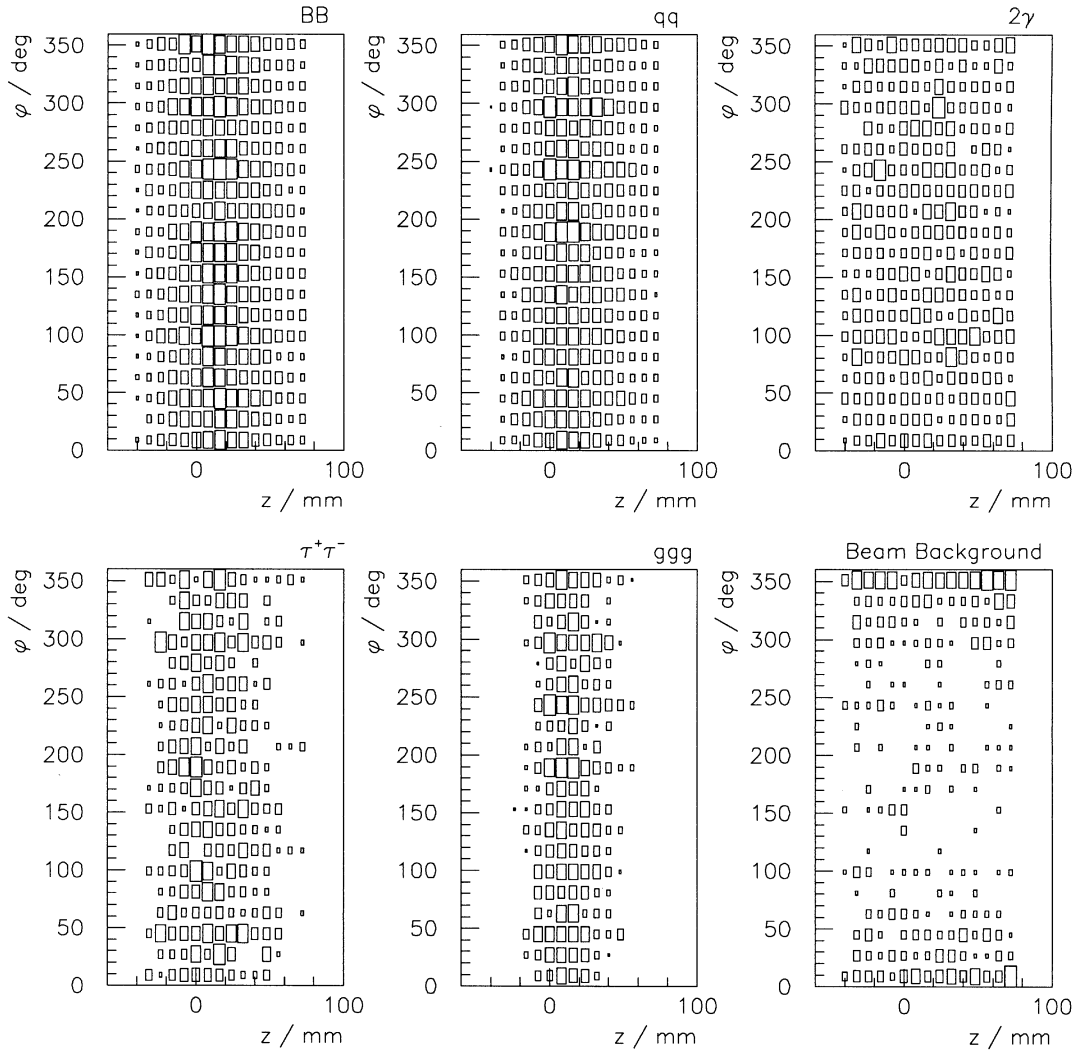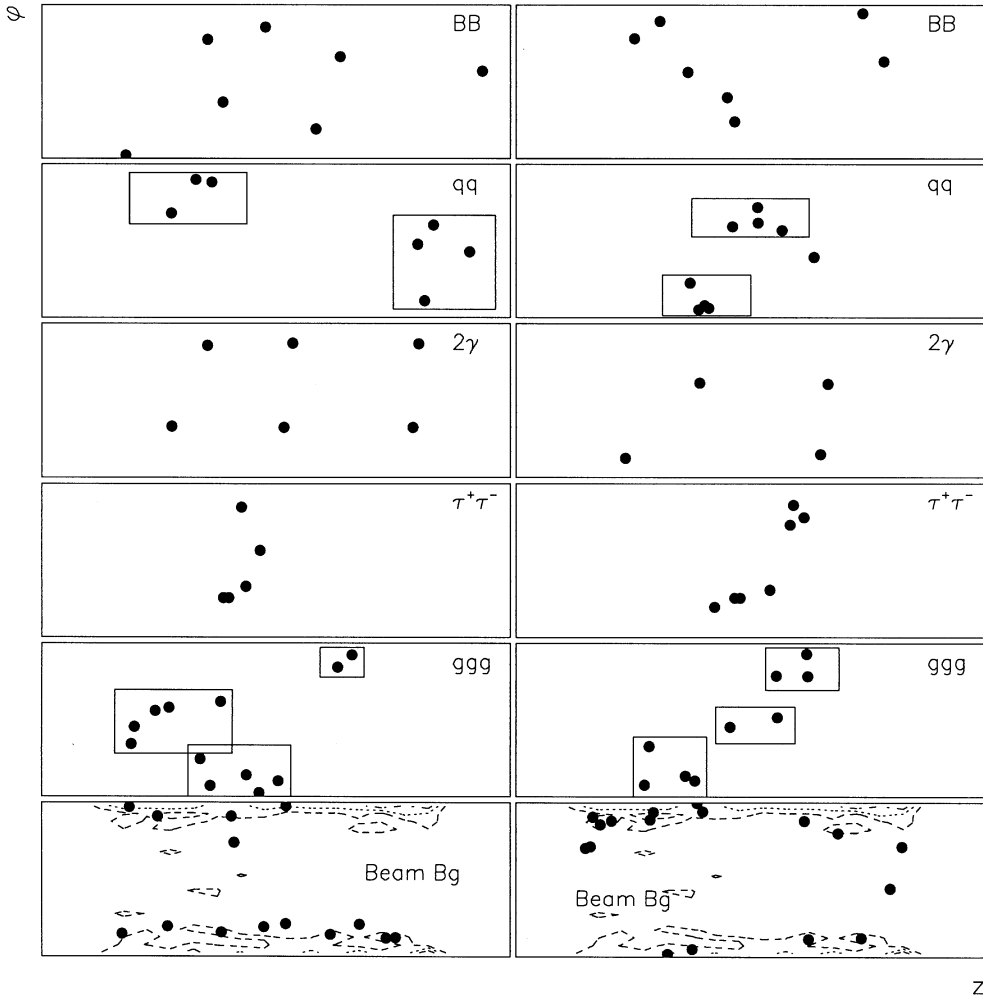
Fig. 3. $z\varphi$ topology for event ensembles of $N = 1000$ events of e^+e^- into different final states at $\sqrt{s} = 10.6$ GeV, z coordinates corresponding to the BELLE SVD inner layer. From left to right, top to bottom: $\Upsilon(4S) \to B^\circ\overline{B}^\circ$, hadron production from continuum $q\overline{q}$, two-photon reactions 2γ, $\tau^+\tau^-$ production by annihilation, three-gluon events (ggg) and beam background.

as $B^\circ(\overline{B}^\circ) \to J/\Psi\,K_s$ will be among the first studies at BELLE. $B^\circ\overline{B}^\circ$ events (and B^+B^- events) show a uniformly spread distribution of hits in the $z\varphi$-plane (cf. Fig. 3 top, Fig. 4 top). Due to the high mass of the produced b quarks a jet structure is generally not observed.

2.2.2. $q\overline{q}$ events

The pair production of a heavy b pair through the $\Upsilon(4S)$ resonance is overlayed by pair production of the light quarks $u\overline{u}$, $d\overline{d}$, $s\overline{s}$ and $c\overline{c}$. The cross section of this hadron production from continuum $\sigma_{qq} = 3.0$ nb is higher than $\sigma_{\Upsilon4s}$. However, the

Fig. 4. $z\varphi$ topology for single events of e^+e^- into different final states at $\sqrt{s} = 10.6\,\text{GeV}$, z coordinates corresponding to the BELLE SVD inner layer. From top to bottom: $\Upsilon(4S) \rightarrow B^\circ \overline{B}^\circ$, hadron production from continuum $q\overline{q}$, two-photon reactions 2γ, $\tau^+\tau^-$ production by annihilation, three-gluon events (ggg) and beam background. $q\overline{q}$ events show a typical two-jet, ggg events a typical three-jet structure (indicated by boxes). In the case of beam background the distribution for 1000 events is added as contour plot (enhancement in the $\varphi \simeq 0$ region).

$q\overline{q}$ events are expected to be separated efficiently from $B^\circ \overline{B}^\circ$ events in offline analysis using thrust and sphericity shower analysis [4]. $q\overline{q}$ events are of theoretical interest as they provide a tool for testing perturbative QCD predictions, even in the next-to-leading order regime [10]. Due to the low mass of the produced light quarks, all ejectiles have relatively high kinetic energy (compared to $b\overline{b}$

production). Thus, $q\overline{q}$ events generally show a back-to-back two-jet signature, with a topology $\Delta\varphi\,[Jet(1),\,Jet(2)\,] \simeq 180°$ (cf. Fig. 4).

2.2.3. ggg events

Gluon events generally show a typical three-jet structure [11]. Three gluon production $e^+e^- \rightarrow$ ggg should be considered here as typical three-jet

example also for other gluon reactions, e.g. gluonic bremsstrahlung $e^+e^- \rightarrow q\bar{q}g$. Experimental data are rare due to the small cross section at $\sqrt{s} = 10.6\,\text{GeV}$, i.e. three-jet events are only 2.9% of all jet events at $\sqrt{s} = 14\,\text{GeV}$ [12]. However, gluon events should be considered here, because BELLE will also perform measurements below the $b\bar{b}$ threshold for background studies, e.g. at the $\Upsilon(1S)$ resonance. Former analyses [13] revealed that the $\Upsilon(1S)$ resonance has an hadronic width nearly saturated by decay into three gluons. Generally, gluon jets show a slightly higher charged multiplicity than quark jets in the order of $\leq 29\%$ [14], depending on the gluon production process.

2.2.4. Two-photon events

The reaction $e^+e^- \rightarrow e^+e^-\gamma^\star\gamma^\star$ is a higher-order QED process which also generates multiparticle final states. Two-photon reaction measurements, especially for hadron production $\pi\pi$, KK, $p\bar{p}$ or $\Lambda\bar{\Lambda}$, are of interest due to implications e.g. for SU(3) flavour symmetry breaking effects [15]. The $\gamma\gamma$ collision axis is nearly coincident with the e^-e^+ beam collision axis, leading to a well balanced event $p_\perp$. Since e^- and e^+ scatter at small angles, they generally leave the experimental area without being detected. As an event generator for hadron production was not available for this study, only $\gamma^\star\gamma^\star \rightarrow \mu^+\mu^-$, including radiation and decay, was taken into account [6]. Multiparticle events with charged multiplicity $N \geq 2$ (42% of the data set) in the inner SVD layer were extracted.

2.2.5. $\tau^+\tau^-$ events

τ particles can be created by the annihilation process $e^+e^- \rightarrow \tau^+\tau^-$. The τ particle decays with a decay length of $c\tau = 91.4\,\mu\text{m}$ into three or more charged prongs with a branching fraction of $> 14\%$, thus generating charged multiparticle states. Generally, the events are not balanced in $p_\perp$, because missing momentum is carried away by neutrinos (from τ decay). The $\tau^+\tau^-$ events are of special interest due to possible examination of CPT invariance [16] by employing triple-momentum correlations of the momenta of the incident beam e^- (or e^+) with the momenta of the outgoing e and μ, (final state of $\tau \rightarrow e/\mu \nu\bar{\nu}$).

2.3. Beam background events

Accelerator beam pipes do not contain perfect vaccum. Both beams e^+ or e^- can interact with a residual gas molecule and deviate from the ideal path. Electromagnetic interaction of first order by Coulomb scattering or second order by Bremsstrahlung is dominant. Hadronic interaction (photo–nuclear reactions) should not be considered here, due to their smaller cross section, i.e. the photo–nuclear interaction length is longer than the radiation length typically by more than two orders of magnitude [17].

The topology can be described by some general properties: (a) high charge hit multiplicity (up to three times higher than physics events' multiplicity, cf. Fig. 6, top left), (b) enhancement in $\varphi \simeq 0$ region (cf. Fig. 1, right) and (c) wide z coordinate spread (cf. Fig. 6, top right). Furthermore, ADC spectra of beam background events differ significantly from physics events due to signal/trigger mismatch, i.e. they are originating from different beam bunches.[2]

The ADC value corresponds to the amplitude of an SVD channel preamplifier output at a given time t_{ADC}. The timing will be adjusted in such a way that in case of a physics event the time t_{ADC} exactly corresponds to (a) the BELLE global trigger time trigger generated by the BELLE central drift chamber or the BELLE time-of-flight system and, at the same time, (b) the maximum of the preamplifier output signal. If $t = 0$ marks signal start of the preamplifier output, this adjusted time is given by $t_{\text{ADC}} = 2\,\mu\text{s}$. In case of a beam beackground event, the trigger occurs at some other time in the time window $0 \leq t_{\text{ADC}} \leq 20\,\mu\text{s}$, thus a randomly chosen amplitude of the rising or declining signal is taken as ADC value. Therefore beam background events show an ADC distribution with nearly exponential shape (cf. Fig. 5A, gray histogram), physics events show a peak corresponding to a charge equivalence of $\sim 15.000\,e^-$ (ADC value ~ 250 in Fig. 5A, white histogram). Table 1 shows cross-section

---

[2] In the KEK-B accelerator beam bunches of $l = 0.4\,\text{cm}$ are spaced by $\Delta l = 59\,\text{cm}$ [2], thus the average time between two bunches is $\tau = 2\,\text{ns}$.
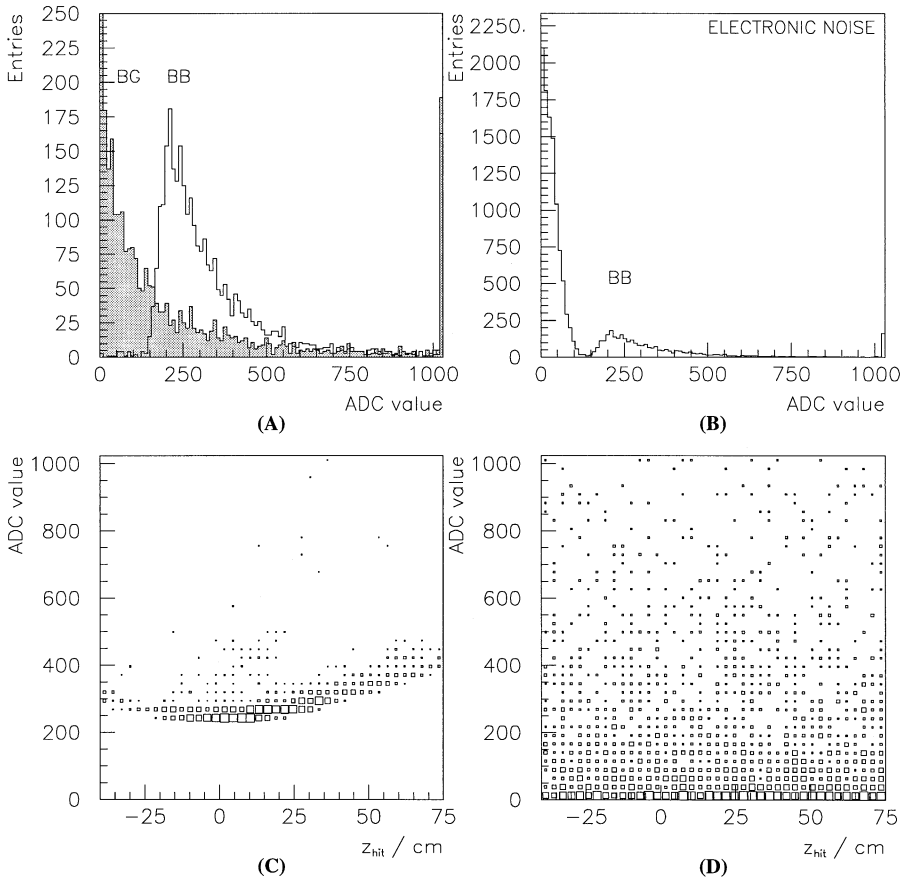
Fig. 5. ADC signatures. (A) Simulated ADC spectra for beam background (gray) and $B^\circ \overline{B}^\circ$ (white). Beam background peaks at channel zero due to signal/trigger mismatch (originating from different beam bunches). (B) Simulated ADC spectrum for $B^\circ \overline{B}^\circ$ including electronic noise (signal threshold at $3 \times \text{RMS}$). (C) Simulated ADC value vs. hit position z_{hit} for $B^\circ \overline{B}^\circ$ events as an example for physics events generally. Charged particles originating from the beam collision point vicinity pass higher z coordinates with higher polar angle, resulting in a higher material path length with higher energy deposition. (D) Simulated ADC value vs. hit position z_{hit} for beam background events show a random pattern (for details see Section 2.3).

σ and expected trigger rates R for different types of R.

Due to geometrical and electronical reasons, SVD hits usually show a cluster structure, i.e. for one charged particle a number $2 \leq n \leq 6$ neighboured strips show significant charge accumulation. The corresponding ADC signals were summed before usage as input for the classification algorithm.

Fig. 5 C and D also show the necessity of an analysis in a higher dimensional input space (in this case $n = 2$) rather than analysing the 1-dim ADC histogram alone. In case of physics events, charged particles originating from the vicinity of beam collision point, pass higher z coordinates with higher polar angle ϑ, resulting in a higher material path length Δl with higher energy deposition ΔE. In a 2-dim histogram ADC vs. z (cf. Fig. 5C) this correlation (ADC value $\sim \Delta E \sim \Delta l \sim \cos(\vartheta)$) is visible as a distribution with curved shape. Beam background events show a different distribution (cf. Fig. 5D), i.e. no dependance $\text{ADC} = f(z)$ is visible.

Table 1
Cross sections σ and expected trigger rates R for different event types at BELLE ($\sqrt{s} = 10.6$ GeV, luminosity $\mathscr{L} = 10^{34}$ cm^{-2} s^{-1}, [1]). The trigger rate value for two-photon events (2γ) assumes a transverse momentum cut $p_t \geq 0.3$ GeV. Values for $\tau^+\tau^-$ include also $\mu^+\mu^-$

Abbreviation	Reaction	σ(nb)	R(Hz)
$B^\circ\bar{B}^\circ$	$\Upsilon(4S) \to B^\circ\bar{B}^\circ$	1.2	12
$q\bar{q}$	$e^+e^- \to q\bar{q}$	2.8	28
ggg	$e^+e^- \to ggg$	< 1	< 1
$\tau^+\tau^-$	$e^+e^- \to \tau^+\tau^-$	1.6	16
2γ	$e^+e^- \to e^+e^-\gamma^\star\gamma^\star$	~ 15	~ 35
Beam background	EM interaction with residual gas	—	≥ 100 Hz

Table 2
Transputer T805 parameters

Clockrate	25 MHz
Integer peak performance	30 MIPS
Floating point peak performance	4.3 MFlops
On chip static RAM	4 kB
Interrupt response time	630 ns

3. Hardware

3.1. The T805 transputer

SGS THOMSON^{tm} inmos^{tm} transputer systems are successfully used for several years due to their ability of parallel data processing. Applications in pattern recognition cover e.g. face recognition [18]. Furthermore, particle physics experiments make use of transputer networks for online purposes such as data acquisition (DAQ) [19] and triggering [20]. In this paper, we present an application of a T805 [21] transputer system for a particle physics pattern recognition task by employing a self-organizing map (SOM) (cf. Section 4.2, [22]). It was shown recently that SOMs can be parallelized successfully on other hardware platforms (e.g. IBM workstation cluster [23]). The basic T805 parameters are listed in Table 2.

The maximum BELLE (all subdetectors) and BELLE SVD DAQ readout frequency is designed to be $R = 500$ Hz [3], being the highest rate used for a particle physics transputer DAQ system so far. Former systems as the DAQ system of the VENUS experiment [24] handled lower trigger rates ($R \leq 15$ Hz). For rates $R > 500$ Hz the transputer should be replaced by more powerful systems as digital signal processors.

In this paper, we study the possible implementation of a beam background rejection algorithm onto the SVD DAQ system. Due to the value of

$R = 500$ Hz in this case, the maximum transputer processing time (maximum time for yes/no decision) per event is limited to $\tau \leq 2$ ms, corresponding to a typical level 3 trigger time.

However, the proposed implementation is flexible and can also be applied for offline analysis purposes, i.e. filtering of data events from storage tapes after the experiment. The event I/O in that case could be performed via the transputer host interface to a workstation (in our case a SUN SPARC-5V running SunOS 4.1.1). This I/O facility was also used for processing the Monte Carlo generated data for this study.

3.2. DAQ system

The proposed DAQ system [19] consists of a T805 network, connected by serial transputer links (bandwidth ≤ 1.8 MB/s). The system serves to read out 81.900 silicon strips.

The DAQ readout sequence is shown in Fig. 2. One data path consists of the four subsequent modules: (a) frontend board for data read, (b) data scanner (DS) for analog-to-digital conversion, (c) slave memory module (SMM) for temporary storage of a partial event and (d) DMA master module (DMM) for local event building. The system consists of 64 modules of (a), 48 modules (b) and (c) each, as well as four DMM modules. Modules (b)–(d) each contain one T805 transputer chip, resulting in a total amount of 103 transputers.

For performing a physics/background separation task, principally one has to examine the topology of the whole event. Therefore the algorithm should be implemented on a DMM or, with modifications of the algorithm, on the SMM network, i.e. using transputer links to examine event parts of a neighboured SMM.

3.3. occam-2 programming

Transputer programming was performed using the programming language occam-2[25]. occam-2 provides standard sequential functionality as well as specific programming abilities for parallel loop execution, e.g. SEQ i = 0 FOR 100 for a sequential loop, PAR i = 0 FOR 100 for a parallel loop. Using such abilities, independent loop iterations in the SOM algorithm can be executed at the same time step.

However, using as much as possible PAR statements does not necessarily increase processing speed. For each loop index i one process is started, containing process operating overhead. Number of processes per transputer CPU is limited, thus transputer program optimization means finding the optimum number of SEQ and PAR statements rather than using PAR in *each* loop alone.

Timing measurements were performed using the data type TIMER. The TIMER default digitization of $\Delta t = 64\,\mu s$ can be changed by using the PRI PAR command to $\Delta t = 1\,\mu s$ according to the scheme

```
TIMER clock :
PRI PAR
  SEQ
    clock ? t1
    ... - event processing
    clock ? t2
  SEQ
    SKIP - necessary dummy
```

In case of a possible offline application (to be discussed in detail in Section 6.6) the event I/O can be performed by usage of the transputer host interface to a workstation with the occam-2 libraries hostio.lib and streamio.lib, e.g. the command ss.gets() for reading an event from standard input.

3.4. Connection to frontend system

The implementation of a possible level 3 trigger into the BELLE trigger system is not studied in detail yet, although the framework has already been prepared. However, the yes/no decision could be transmitted as a logical signal to the LeCroy 1810 module [26] which is part of the BELLE FASTBUS data acquisition. This module can, according to FASTBUS specifications, distribute a *fast clear signal*, which stops the started analog-to-digital data conversion and resets all modules within $\tau \leq 500\,ns$. Typical cable delays are in the order of $\tau \leq 100\,ns$ for cable lengths $L \leq 20\,m$. Thus, further signal processing for usage in the BELLE trigger scheme would not be time critical.

4. Algorithm

4.1. Input

A schematic overview of the algorithm is shown in Fig. 8. The basic principle of the algorithm is a topology classifier in the unfolded $r\varphi$-plane of the inner SVD layer, followed by a mathematical tool for an analysis in a 5-dim input space, namely a Self-Organizing Map (SOM) (cf. Section 4.2).

Each of the physics event types (cf. Section 2.2) as well as the beam background events (cf. Section 2.3) show more or less characteristic topologies in the $r\varphi$-plane. Topologies for event ensembles are shown in Fig. 3, topologies for single events in Fig. 4. The characteristics of the topologies are used for preparation of five variables which are fed into an SOM. The five input parameters x_i, $i = 1,\dots,5$ are listed in Table 3, displayed in Fig. 6 and will be discussed in detail here.

Input parameter 1: Physics events usually generate signals on 60–80 strips in the first SVD layer,

Table 3
Input and output parameters for the self-organizing map. For details see Section 4.1

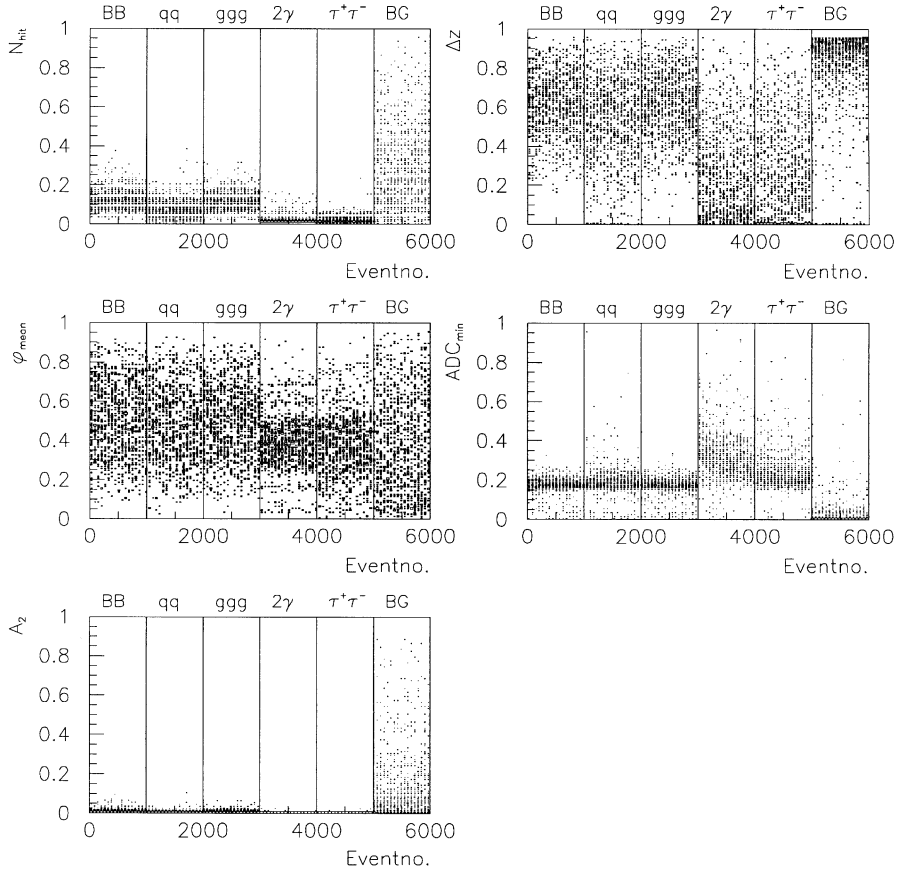| Parameter	Definition	Transformation
Input		
N_{hit}	Number of charged hits	$1/N_{hit}$
Δz	z spread $z_{max} - z_{min}$	$1 - \Delta z$
φ_{mean}	φ mean value $(\sum_{i=1}^{N_{hit}} \varphi_i)/n$	—
ADC_{min}	Minimum ADC value	$ADC_{min} \times 2$(streched)
A_2	Two-point correlation value	$1/A_2$
Output		
d_{min}	Minimum 5-dim Euclidean distance	

Fig. 6. SOM input parameters (normalized) vs. event number. From top to bottom, from left to right: Number of charged hits N_{hit}, z spread $\Delta z = z_{max} - z_{min}$, φ_{mean}, ADC_{min}, two-point correlation value A_2. For details see Section 4.1.

beam background events up to a maximum of $\simeq 300$ strips. Thus, a high value N_{hit} is a good indicator for the presence of a beam background event (cf. Fig. 6, top left).

Input parameter 2: Among all charged hits the minimum and maximum z coordinates z_{min} and z_{max} are searched. Beam background events generally show a larger $\Delta z = z_{max} - z_{min}$ value than physics events, because the particle tracks are not originating from the beam collision point (cf. Fig. 6, top right).

Input parameter 3: Beam background events show a strong enhancement in the $\varphi \simeq 0$ region (cf. Fig. 1, right). Thus, the azimutal angle average value $\Delta\varphi = \sum_{i=1}^{N} hit\varphi_i \simeq 0$ provides

a good classification tool (cf. Fig. 6, middle left).[3]

Input parameter 4: The minimum ADC value ADC_{min} (cf. Fig. 6, middle right) is a good indicator for a beam background event due to signal/ trigger mismatching (particles originating from different beam bunches). Details are described in Section 2.3.

Input parameter 5: Two-point correlation functions are widely used in astronomy [27] to describe

---

[3] The exact calculation is $\Delta\varphi = \sum_{i=1}^{N} hit\ \varphi'_i \simeq 0$ with the two cases (a) $\varphi' = \varphi$ for $\varphi < 180°$ and (b) $\varphi' = (360° - \varphi)$ for $\varphi > 180°$ (symmetrization on a cylindrical surface).
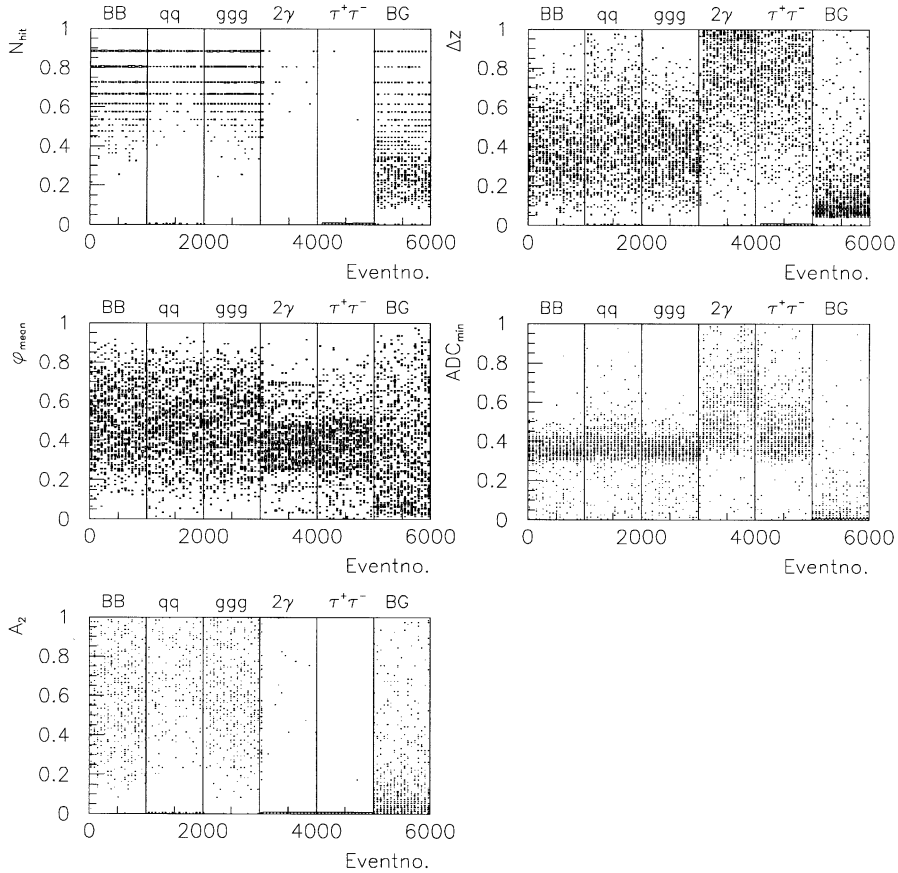
Fig. 7. Mathematically transformed SOM input parameters (normalized) vs. event number. Transformations are described in Table 3. For sequence see Fig. 6. Beam background events should cluster at the bottom ($x_i \simeq 0$), physics events at the top ($x_i \simeq 1$). For details see Section 4.1.

galaxy clustering. Recently, it was shown that they are also applicable to cluster analysis in particle physics data analysis [28]. All distances between data points (in our case distances between charged hit coordinates in the $z\varphi$-plane) are calculated pairwise and filled into a histogram. If points are clustered in certain regions, the distances and the mean value of the histogram, defined as A_2, are small. If points are not clustered, A_2 is large. In our case beam background events often do not show any small scale clustering as it is given in cases of jet production e.g. for $q\bar{q}$ or ggg, resulting in large A_2 values (cf. Fig. 6, bottom). However, as already mentioned, they show a clustering in the $\varphi \simeq 0$

region. Thus, the combined information of the input parameters $x_3 = \varphi_{\mathrm{mean}}$ and $x_5 = A_2$ is important for classification.

All input parameters are normalized to a range [0,1] in order not to allow decision priority to a single parameter (by dominating the 5-dim Euclidean distance with the value of the single parameter's 1-dim distance). Moreover, all parameters x_i were transformed by simple mathematical transformations (cf. Table 3, Fig. 7) in order to guarantee maximum classification efficiency. The aim of the transformations was to cluster beam background events around the $x_i \simeq 0$ region, physics events around the $x_i \simeq 1$ region.

4.2. Training

Generally, neural networks must be *trained* to perform a certain classification task, a *recall* phase for data classification follows after *training*. Algorithms must be distinguished into two related groups: (a) *supervized* models to be trained with Monte Carlo simulated data holding a complete input–output information and (b) *unsupervized* models using only input information. In the latter case the network is trained directly with unbiased experimental data and performs a search for clusters in a given n-dimensional input space.

The SOM [22] is the most popular *unsupervized* neural network algorithm, used in many applications as a mathematical tool for an analysis in an n-dim input parameter space. A SOM is based on units called *nodes* which can be defined as points in the n-dim input space representing a complete cell, also called Voronoj cell, around it. An n-dim cluster in the input data is filled with cells and thus quantized. The cells can be compared to bins in a 1-dim histogram. The number of nodes N_{node} thus determines the accuracy of binning. In the recall phase for an n-dim event vector the Euclidean distances to all nodes are calculated. The minimum Euclidean distance d_{min} provides a proper tool for classification.

Mathematically, the SOM can be regarded as a nonlinear function $f(x) = O$, where x is an n-dim vector, O is a scalar output variable. It was proven that f is *topology preserving*, i.e. even complex cluster structures (e.g. an n-dim torus) can be mapped [22].

If a SOM is trained on several clusters, the clusters must be identified in the first step. In the second step, d_{min} determines the corresponding cluster and hence the classification. In this paper, we train the SOM only on one cluster, i.e. the cluster given by 5-dim beam background data. A beam background event should have a small d_{min}, whereas a physics event a large d_{min} in the recall phase. As we used Monte Carlo data for this study, the identification of beam background data is fixed. In the experiment, pure beam background data could be recorded by running the accelerator in non-collision mode, e.g. either with e^+ or e^- beam only.

However, our SOM algorithm (Fig. 8) differs from the most simplest approach by employing two features which were introduced into self-organizing neural networks rather recently: *growing* and *gravitation*.

4.2.1. Growing

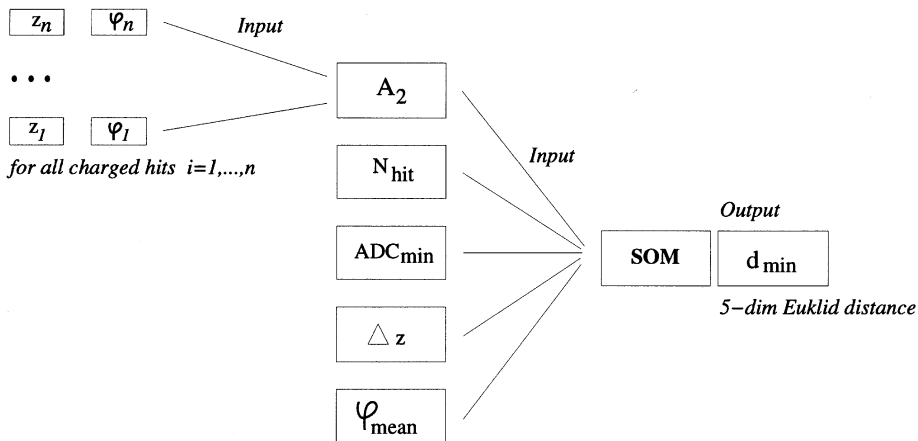The original Kohonen algorithm is based on a fixed number of nodes N_{node}, not changed during

Fig. 8. Algorithm scheme. All charged hits in the $z\varphi$-plane are used for the calculation of the two-point correlation value A_2. The SOM analysis is performed in a 5-dim input space (cf. Section 4.1). The SOM output is a minimum Euclidean distance d_{min}, serving for an yes/no decision according to a cut value Δd_{min}^{cut}.
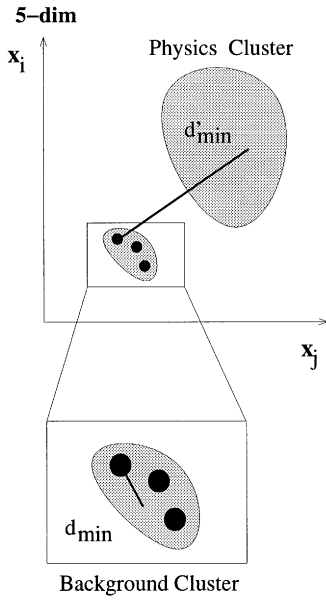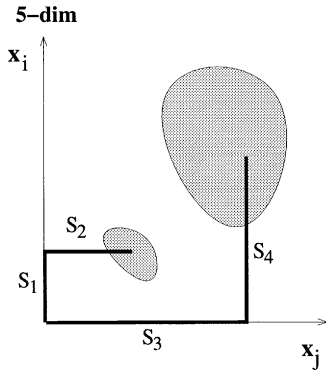
SOM Classifier

Alternative Linear Classifier

Fig. 9. Classification principle. Top: SOM classification. In the 5-dim input space the beam background events form a cluster near the origin $x = 0$ (cf. Fig. 7), the physics events far from the origin. For an event e the Euclidean distance $d(e,n)$ to all nodes n is calculated and the minimum distance d_{min} is chosen. The distance d_{min} is small for beam background events (zoomed cluster) and large for physics events. Bottom: Linear classifier. The sum of the input variables S also provides a classification tool (cf. Section 6.1). In this case, $S = S_1 + S_2$ for a beam background event is smaller than $S' = S_3 + S_4$ for a physics event.

the training phase. The number of nodes is a free parameter, to be adjusted by the user. Recently, SOM algorithms, extensions with variable

N_{node} were published (Growing Cell Structures (GCS) [29], Growing Neural Gas (GNG) [30]). Creation and destruction of nodes, according to the actual SOM status, is a part of the algorithm. In our application, we use a simplified GCS ansatz. In the training phase, with the first processed event vector x ($e = 1$) a node vector $x(n)$ is created, with exactly the same 5-dim coordinates (Fig. 9).

For each event $e = 1, \ldots, N_{event}$ with the event vector $x(e)$ the Euclidean distances to all formerly created nodes $n = 1, \ldots, N_{node}$ are calculated:

$$d(e,n) = x(e) - x(n) \sqrt{\sum_i^{N_{dim}} (x_i(e) - x_i(n))^2}. \qquad (1)$$

All distances $d(e,n)$ are compared to a predefined cut value[4] d_{cut}, corresponding to the radius of the Voronoj cell. In case of $d(e,n) > d_{cut}$, a new node with the event vector coordinates is created. If $d(e,n) < d_{cut}$ is valid for all nodes n, then the node with the smallest Euclidean distance $d_{min} = \min[d(e,n)]$ is searched among all nodes. In SOM algorithms this node usually is called *winner node*. As will be described in Section 4.2.2, each node has a *mass* by definition, due to the gravitational part of the algorithm. The mass of the winner node is increased at this step. Thus, the mass of the node is correlated to the number of events in the Voronoj cell, and keeps the information, the node is very important for the classification process.

4.2.2. Gravitation

To achieve a proper mapping of the n-dim cluster structure, a mechanism must be provided to move the nodes in the input space. In the original Kohonen SOM [22] all node vectors $x(n)$ are moved into the direction of the actually processed event vector $x(e)$ (with free parameters ε and δ):

$$x(n) = x(n) + \varepsilon(t) \cdot f(\delta,t) \cdot (x(e) - x(n)). \qquad (2)$$

However, in a former study [31] it turned out that gravitational interaction between the nodes also provides a feasible mechanism for movement. In the gravitational SOM, each node is assigned with

---

[4] Typical cut values are $0.01 \leq d_{cut} \leq 0.15$. The setting of this value automatically determines the final number of nodes.

a mass value m, simply defined as an integer number and initialized for each node to $m = 1$. As described in Section 4.2.1, the mass is increased $m = m + 1$ for winner nodes during the training. Thus, the mass of a node contains information about the event density in the Voronoj cell of a node and hence, the node is vert important for the classification process.

Concerning the algorithm itself, there are two main motivations for its usage instead of the original Kohonen algorithm:

1. In case of the original Kohonen algorithm, the optimization of the parameters ε and δ as well as the time dependance $\varepsilon = f_1(t)$ and $\delta = f_2(t)$ are very time consuming; general rules do not exist. The gravitational algorithm employs only one parameter α, corresponding to a gravitational force strength, which is not time dependent. Moreover, the parameter α is correlated to the typical n-dim cluster size [31] and can therefore be easily estimated.

2. The gravitational algorithm is less CPU time consuming than the original Kohonen algorithm. During the training phase the original SOM updates the node vectors with each iteration, i.e. with the processing of each single input event. The training phase of the gravitational SOM is divided into two parts. In the first step all events are processed once, the nodes are created and the node masses are updated. In the second step, after event processing, the gravitational contraction of the nodes is performed. The number of iterations corresponds to $\mathcal{O}(N_{\text{events}} \times N_{\text{node}})$ for the original SOM and to $\mathcal{O}(N_{\text{node}}^2)$ for the gravitational SOM. It should be emphasized that gravitation is only part of the training, not part of the analysis process.

In the gravitational SOM, for each node $n = 1, \ldots, N_{\text{node}}$ the resulting gravity vector $\mathbf{g}(n)$ is calculated and the corresponding node vector $\mathbf{x}(n)$ updated according:

$$d(n',n) = \mathbf{x}(n') - \mathbf{x}(n), \tag{3}$$

$$\mathbf{g}(n) = S \cdot d(n',n), \tag{4}$$

$$\mathbf{x}'(n) = \mathbf{x}(n) + \mathbf{g}(n). \tag{5}$$

The ansatz for the gravitational strength S, as originally proposed in [32], is given by

$$S = \frac{\alpha}{\alpha + d} \cdot m(n), \tag{6}$$

where the gravitational force strength α is given by the average cluster size (in our case $\alpha \simeq 0.1$). However, for purposes of faster transputer processing this function was replaced by $S' = (1 - d)^{c_1}$ with a similar shape and comparable classification results (typical values $8 \leq c_1 \leq 32$).

The effect of the gravitational algorithm is shown in Fig. 10 for a 2-dim projection of the 5-dim input space. Fig. 10C shows the node distribution before, Fig. 10D shows the much better node mapping after the gravitational contraction. Training is performed with beam background data only. The corresponding beam background data distribution that should be mapped is shown in Fig. 10A.

4.3. Output

In the SOM analysis phase, for each event e to be classified after training, again the Euclidean distance $d(e,n)$ to all nodes n is calculated according to Eq. (1). As described in Section 4.2, the minimum Euclidean distance $d_{\text{min}} = \min[d(e,n)]$ is a proper tool for classification, i.e. d_{min} should be small for beam background events and large for physics events. The classification principle is graphically shown in Fig. 9 (top).

Fig. 11A shows the value of d_{min} as a function of the event number, 1000 events representing the data set for one reaction type. Fig. 11B shows 1-dim histograms for d_{min} for the cases of beam background and physics events. Final classification is done by defining a cut value $d_{\text{min}}^{\text{cut}}$ which is given by the position of the minimum in an histogram containing the d_{min} values for all beam background and physics events (corresponding to the (SOM histogram) grey and white histograms, respectively, in Fig. 11B). In the analysis, a value $d_{\text{min}} \leq d_{\text{min}}^{\text{cut}}$ leads to classification as beam background and hence event rejection, $d_{\text{min}} > d_{\text{min}}^{\text{cut}}$ to classification as physics event which must be kept.
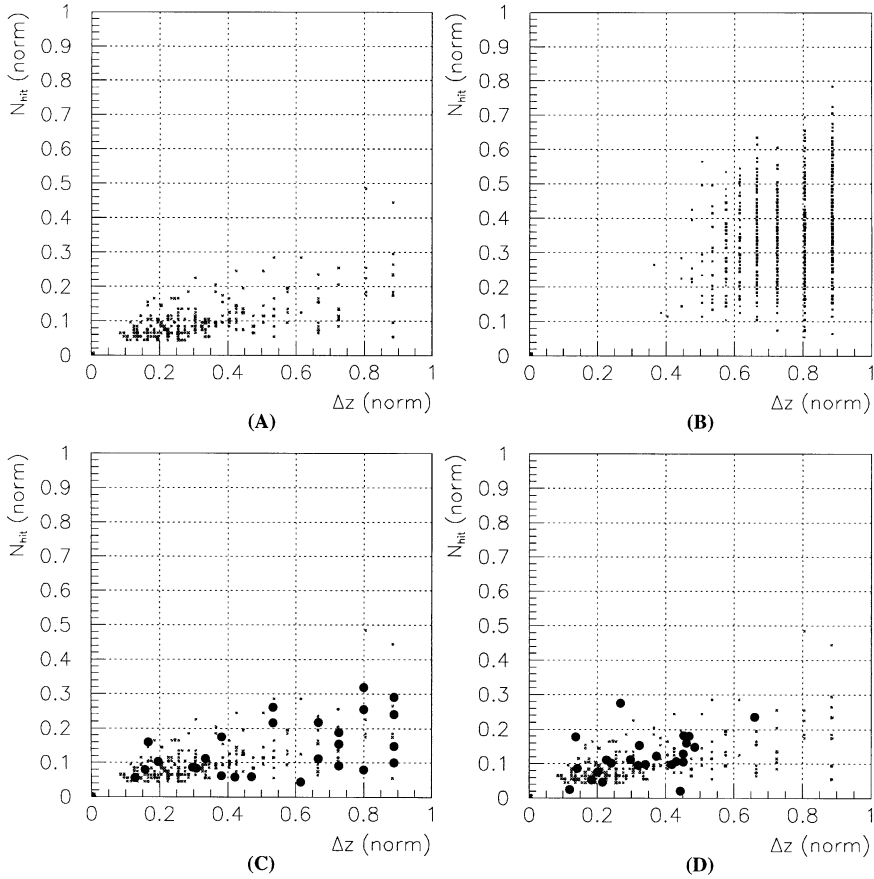
Fig. 10. Gravitational algorithm. 2-dim projection (N_{hit} vs. Δz) of the 5-dim input space for (A) beam background data, (B) physics data, (C) node positions before and (D) node positions after gravitational phase. After the gravitational contraction the nodes map the beam background event distribution shown in (A) properly.

5. Results

5.1. Classification efficiency

The classification performance can be evaluated with two parameters: (a) background rejection fraction F_{BG} and (b) physics pass fraction F_{PHYS}. Both F_{BG} and F_{PHYS} should be as high as possible. Fig. 12 shows the final results as a function of the necessary CPU event processing time on the transputer. Within a decision time of $\tau = 2$ ms values of $F_{BG} = 75.0\%$ and $F_{PHYS} = 96.9\%$ were achieved. Detailed values of F_{PHYS} for the five different event types are listed in Table 4.

As the fluctuation for F_{BG} in Fig. 12 is smaller than the estimated error (cf. Section 5.3), it can be concluded that SOM convergence is achieved. The high F_{BG} data point at $\tau = 0.6$ ms CPU time is assumed to be a local classification minimum, not suitable for the final trigger application, because at the same time less physics events pass the trigger (cf. Fig. 12, bottom).

5.2. Timing measurements

Realtime timing measurements were performed using the occam-2 data type TIMER (cf. Section 3.3). In the training phase, due to pairwise
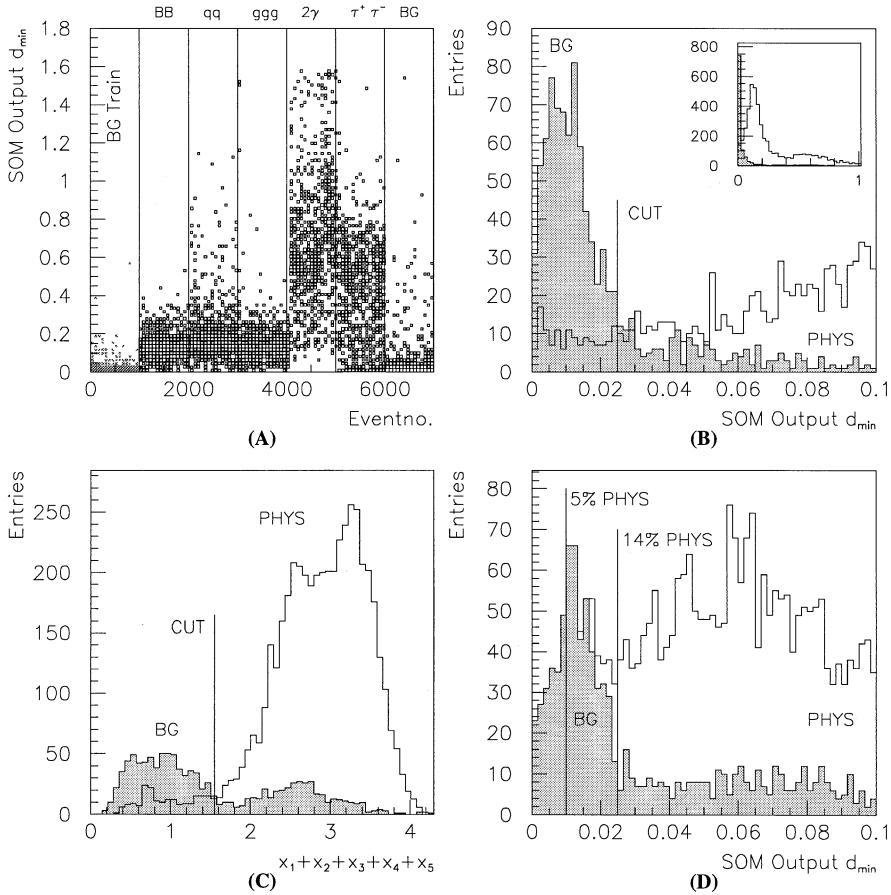
Fig. 11. SOM output. (A) SOM output vs. event number e. The event number serves as an indicator for the reaction type. Events with $e \leq 1000$ are training events (beam background only, grey). (B) Histogram of SOM output d_{min} for beam background events (grey) and physics events (white) for full output value range (*small histogram*) and zoomed cut region (*large histogram*). The cut position d_{min}^{cut} is displayed. Events with $d_{min} \leq d_{min}^{cut}$ are classified as beam background, events with $d_{min} > d_{min}^{cut}$ as physics events. (C) Linear classifier as an alternative method. The input sum $S = x_1 + x_2 + x_3 + x_4 + x_5$ is histogrammed. Beam background rejection fraction (68.3% for a cut position $S_{cut} = 1.55$) is smaller than in the case of SOM classification (75.0%, cf. Section 6.1). (D) SOM output in case of electronic noise (signal threshold set to 3RMS). For details see Section 6.5.

node–node gravity calculations the CPU time t is a quadratic function of the number of nodes N_{node} (cf. Section 4.2). All measurements could be fitted ($\chi^2 < 18 \times 10^{-5}$) to

$$t[\text{ms}] = 0.07\, N_{node}^2 - 0.08\, n - 0.15. \qquad (7)$$

Thus, the algorithm needs $t = 6$ ms per event in case of 10 nodes, $t = 680$ ms for 100 nodes. This total event processing time includes both the mass update phase (event iteration) and gravi-

tational contraction phase (node iteration) (cf. Section 4.2.2).

In the analysis phase gravitation is not needed. Therefore $t = f(N_{node})$ is a linear function which could be fitted ($\chi^2 < 2 \times 10^{-5}$) to

$$t[\text{ms}] = 0.022\, N_{node} - 0.024. \qquad (8)$$

10 nodes need $t = 0.18$ ms per event, 100 nodes need $t = 1.95$ ms. Therefore, a binning in the 5-dim input space using 100 nodes is the accuracy limit to
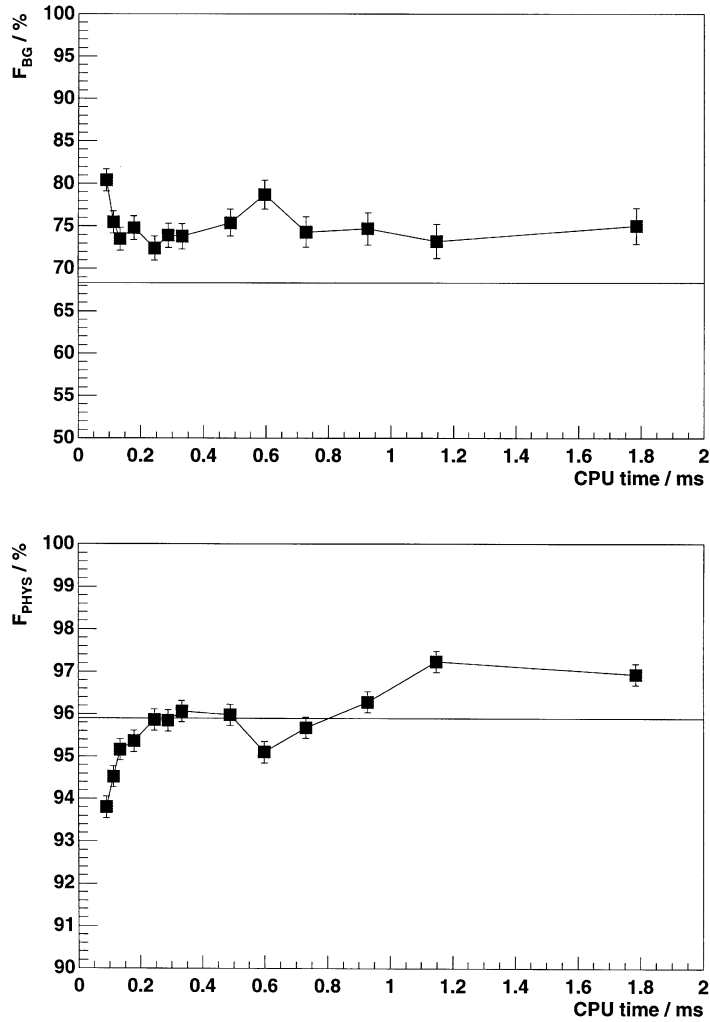
Fig. 12. Results. Top: Beam background rejection fraction F_{BG} vs. CPU decision time. Bottom: Physics trigger passing fraction F_{PHYS} vs. CPU decision time. For details see Section 5. The result for the alternative linear classifier (cf. Section 6.1) is given by a solid line in both figures.

keep the decision time restriction of $\tau \leq 2\,ms$ ($R = 500\,Hz$ readout rate).

5.3. Error estimation

The classification error is given by the error in the determination of the cut position d_{min}^{cut} as displayed in Fig. 11B. As described in Section 4.3 the value of d_{min}^{cut} is determined by the position of the minimum in a histogram, filled with the d_{min} values

for all beam background and physics events. This histogram has a certain binsize, thus the classification error can be estimated by counting the different classification fractions by moving d_{min}^{cut} by $\Delta d = \pm 1$ bin width. Typically, error for beam background is given by $\sim 2\%$ (due to the steep rise of the distribution against $d_{min} = 0$) and by $\sim 0.25\%$ for physics events (flat distribution in the cut region). Errors have been included as error bars in Fig. 12.

Table 4
Physics pass fractions F_{PHYS} for the five different physics event types (cf. Section 2.2)

Event type	F_{PHYS}
$B^\circ \bar{B}^\circ$	97.1%
$q\bar{q}$	97.4%
ggg	97.3%
2γ	> 99.5%
$\tau^+\tau^-$	93.0%
Total	96.9%

Table 5
Comparison of SOM and linear classifier $S = x_1 + x_2 + x_3 + x_4 + x_5$. For mentioned values of O_1, O_2 (cf. Fig. 11B and C). For details see Section 6.1

	Beam BG rejection	Physics pass	Beam BG output O_1	Physics output O_2
SOM	75.0%	96.9%	~ 0.01	~ 0.2
S	68.3%	95.9%	~ 1.0	~ 3.0

6. Discussion

6.1. Alternative linear classifier

Fig. 7 shows that all transformed input parameters x_i have small values $x_i \simeq 0$ for beam background and large values $x_i \simeq 1$ for physics events. A simple alternative method to an SOM could be classification by a sum $S = x_1 + x_2 + x_3 + x_4 + x_5$. The classification principle is graphically shown in Fig. 9, bottom. The classifier S is histogrammed in Fig. 11C. As in case of the SOM output (cf. Section 4.3), a cut value S_{cut} is defined by the minimum position of the sum of the outputs of beam background and physics events (e.g. $S_{cut} = 1.55$ in Fig. 11C). Each event can be classified as beam background for $S \leq S_{cut}$, as physics event for $S > S_{cut}$. The results are worse than in case of SOM classification: 68.3% beam background rejection (75.0% for SOM) and 95.9% physics pass fraction (96.9% for SOM). The reason is given by a worse yes/no signal fraction of the typical beam background output O_1 to the typical physics event output O_2 (cf. Table 5): $O_1/O_2 \simeq 0.01/0.2 = 20$ (cf. Fig. 11B) in case of SOM classification, $O_1/O_2 \simeq = 1.0/3.0 = 3$ (cf. Fig. 11C) in case of linear classification.

6.2. Comparison to other trigger systems

The B meson experiment BABAR at SLAC [33] also plans to use a beam background rejection implementation on a level 3 trigger, using information of the BABAR silicon vertex tracker detector SVT. The input trigger rate is $R = 2000$ Hz (decision time $\tau = 0.5$ ms). The algorithm is based on vertexing, i.e. charged particle track extrapolation onto the z axis, followed by a multiplicity counting in the beam collision point vicinity. The design beam background rejection fraction is > 95%. As in our case within a decision time of $\tau = 0.5$ ms an efficiency of 73.5% was achieved (see data point at $\tau = 0.5$ ms in Fig. 12, top), the BABAR algorithm seems to be a more effective approach. A physics event rejection fraction is not published until now.

The experiment H1 at DESY uses a level 2 trigger [34] for tagging of semileptonic B decays, based upon a feed-forward neural network (FFNN, cf. Section 6.3) algorithm. The decision time is $t = 18\,\mu s$, using the neurocomputer Adaptive Solutions CNAPS. This algorithm was not considered in our case due to the necessity of supervized training by Monte Carlo data with a given input–output information (cf. Section 6.3). Another FFNN trigger system for rejection of background events is proposed for H1, using a fast analog chip [35]. Due to a very short decision time of $\tau = 96$ ns the system could be used for a level 1 trigger decision.

6.3. Comparison to feed-forward neural networks

The popular feed-forward backpropagation neural network (FFNN) algorithm has been used for several applications in particle physics [36]. In a recent comparison [37] an FFNN and a growing SOM performed equally well in case of a particle physics classification task (supervized training mode).

However, an FFNN can only be trained in supervized mode, using a Monte Carlo simulated data

with given input–output correlations. The Monte Carlo model must describe the experimental reality almost perfectly. In our case, the beam background simulation program applies kinematical cuts for the particle event generation, such as a photon energy cut for Bremsstrahlung and a scattering angle cut for Coulomb scattering [7]. Hence, some real event types with certain topologies are not part of the simulated data set. For this reason the choice of an *unsupervized* neural network algorithm as the SOM was essential to avoid any biased classification.

Moreover, understanding the decision process of an FFNN is a difficult task, pointed out in a recent review [38]. Usually an FFNN has a multilayered structure of a input layer, one or two "hidden" layers and an output layer of nodes. For an understanding of the decision process (i.e. why is this event classified as beam background?) an analysis of the hidden layer node values must be performed, using time consuming statistical techniques as e.g. a Principal Components Analysis (PCA). In case of an SOM, the decision process can easily be understood, because it is given by the minimum Euclidean distance only.

It should be mentioned that in case of an application of the presented SOM algorithm in an envisaged BELLE experiment, the training should be redone using the experimental data instead of the simulated data used for this study. In case of an FFNN application this re-training is usually omitted, because a perfect Monte Carlo data set is assumed. However, as the SOM training on the T805 ($N = 1000$ beam background events) takes only about $\tau \simeq 11$ min (cf. Section 5.2), any problematic CPU time consumption by re-training is not expected.

6.4. Comparison to jet algorithms

Any online implementation of an offline jet algorithm [39], replacing the two-point correlation A_2, was not considered to be appropriate due to two reasons.
1. Variables as thrust, sphericity or planarity are based on momentum information ($p_\perp$ or $p_\parallel$) which is not available in the SVD data stream.

2. Standard analysis is performed in cone regions around a somehow defined jet axis (e.g. thrust axis). In comparison, the two-point correlation A_2 does not need an axis nor cone definition and is therefore a simpler and faster approach.

6.5. Electronic noise

SVD hits which are generated by electronic noise effects rather than by charged particles can be simulated by a GEANT based program, prepared by Ozaki [40]. However, for reasons of easier data handling on the transputer we generated electronic noise hits by parametrization of results of the GEANT simulation.

For an event with N_{hit} hits (data set described in Section 2.2) further N_{hit}^{noise} electronic noise hits were added. The number N_{hit}^{noise} depends on a signal threshold setting S_{thr}, the threshold being measured in terms of the R.M.S. of the simulated electronic noise [40] (a typical experimental value will be $S_{thr} = 3 \times$ R.M.S.).

Total SVD hit rates $N_{hit} + N_{hit}^{noise}$ were determined using the GEANT program and are listed in Table 6 (averaged for $N = 10.000$ events).

Each electronic noise hit was assigned with a randomly generated z coordinate and ϕ coordinate (uniform probability distribution).

According to [40] the ADC value distribution was parametrized with a gaussian probability

Table 6
Effect of electronic noise. Total number of SVD hits $N_{hit} + N_{hit}^{noise}$, averaged for $N = 10.000$ events, for different threshold settings S_{thr} (threshold measured in terms of the electronic noise RMS)

S_{thr}	$N_{hit} + N_{hit}^{noise}$
With electronic noise	
1RMS	29 192
2RMS	5016
3RMS	574
4RMS	283
5RMS	262
6RMS	246
Without electronic noise	
3RMS	289

distribution

$$p = \exp\left(-\frac{1}{2} \cdot \left(\frac{\text{ADC} - c_2}{c_3} \right)^2 \right). \qquad (9)$$

The parameters c_2 and c_3 are based on fits to the simulation results. The ADC spectrum is shown in Fig. 5B. During summation of the ADC values of neighboured stripes, in case of electronic noise, one must take into account that the values can principally be negative [40].

The electronic noise has several consequences for the SOM input parameters (cf. Section 4.1):

1. Due to the random distribution of the electronic noise hit location the topology input parameters Δz, φ_{mean} and A_2 are affected in a negative way, i.e. the $r\varphi$ topology signature of the event is "washed out".
2. The number of charged hits, now given by $N_{\text{hit}} + N_{\text{hit}}^{\text{noise}}$, is still an important parameter for classification, because $N_{\text{hit}}^{\text{noise}}$ has the same value for both beam background and physics event types. Thus, any difference in N_{hit} can still be used for classification.
3. The ADC values for beam background events (cf. Fig. 5A) and for electronic noise hits (cf. Fig. 5B) are very similar, i.e. in both cases low ADC values near zero value are dominating. Therefore the presence of hits with low ADC values is not a good indicator for a beam background event anymore. The only difference is given by the different probability distributions, i.e. nearly exponential shape in case of beam background, gaussian shape in case of electronic noise. However, the identification is a difficult task for an SOM as well as for each alternative classifier. For a better classification the input parameters can be filtered during the training phase by applying an ADC value cut, i.e. only hits with an $\text{ADC} > \text{ADC}_{\text{CUT}}$ are used.

The SOM output in case of electronic noise is shown in Fig. 11D. If the same output cut position is applied as in Fig. 11B (case without electronic noise), a fraction of $F_{\text{PHYS}} = 14\%$ would be rejected. If one wants to keep the maximum rejection value of $F_{\text{PHYS}} = 5\%$, the cut position must be moved to lower SOM output values. In the latter case, using an ADC cut position during training of

$\text{ADC}_{\text{CUT}} = 80$, the final results were achieved (S_{thr} denoting the signal threshold):

$$F_{\text{BG}} = 40.3\% \text{ for } S_{\text{thr}} = 3 \times \text{RMS},$$

$$F_{\text{BG}} = 30.2\% \text{ for } S_{\text{thr}} = 2.5 \times \text{RMS}.$$

Thus, a beam background rate R can be reduced by a factor of 2.5 and 3.3, respectively, in the presence of electronic noise.

6.6. Offline usage

If the KEK-B design values [2] for beampipe pressure were achieved from the beginning of accelerator commissioning, there would not be any need for any *online* beam background rejection system. Although the classification system was designed for online application in the SVD DAQ system, *offline* usage for data filtering purposes is also possible by employing the transputer host interface to a SUN workstation (cf. Section 3.1).

In offline usage, there is no time limit for classification, as it is given in the online case by the maximum DAQ readout rate ($\tau = 2\,\text{ms}$, $R = 500\,\text{Hz}$, cf. Section 1). However, Fig. 12 shows that any improvement of classification would not be achieved by extension of the decision time to values $\tau \geq 2\,\text{ms}$, i.e. both F_{BG} and F_{PHYS} are converging.

It should also be stressed that, due to its simple structure, the algorithm could easily be implemented onto other fast hardware platforms as Motorola 88100 chips (part of the BELLE event building scheme). However, studies revealed that any implementation to standard UNIX workstation systems as SUN SPARC-10 would increase the event processing time to values $\tau > 40\,\text{ms}$ due to lack of parallelization features.

7. Possible future improvements

It should be discussed if any extension of the presented algorithm can lead to higher values of both F_{BG} and F_{PHYS}. Two possible extensions should be mentioned:

1. The SOM algorithm is based on 2-dim topological $z\varphi$-information. Thus, an improvement can

be expected from employing 3-dim information, extracted from a tracking analysis. This approach is used by the BABAR level 3 trigger system (cf. Section 6.2). However, as tracking is a CPU time consuming task, probably another fast hardware platform (e.g. SHARC ADSP 21062 digital signal processors as part of the extension of the SVD DAQ system) must replace the transputer in this case.

2. As can be seen from the contour plot in Fig. 4 (bottom), beam background events do not show a uniform multiplicity distribution $M = f(z)$ as a function of the z-ccordinate, i.e. M has a minimum around $z = 0$. This behaviour is due to the beam pipe geometry and can also be observed in the distribution of track vertices of spent electrons in the BELLE central drift chamber [17]. For this study, only the z-coordinate spread Δz was used as input for the SOM.

Deviations from a uniform $M = f(z)$ distribution are also proposed for a possible FFNN level 1 trigger at the H1 experiment (cf. Section 6.2, [35]) without any further topological information. High background rejection efficiencies $F \geq 80\%$ for simulated H1 data also imply, to some extent, a possible successful application for beam beackground rejection at BELLE.

8. Summary

We studied a possible transputer implementation of a beam background rejection algorithm for the BELLE SVD with GEANT simulated data. The successful algorithm was based upon a growing, gravitational self-organizing map algorithm (SOM, $N_{\text{node}} = 100$, 5 topological input parameters). A high fraction of 75.0% of beam background events can be rejected, 96.9% of physics events ($B^\circ \overline{B}^\circ$, $q\overline{q}$, ggg, 2γ, $\tau^+\tau^-$) pass the classification. The method was compared to linear classification ansatz, rejecting 6.7% less beam background events than the SOM. Due to a decision time of only $\tau \leq 2$ ms, the system could be implemented for on-line use as level 3 trigger. Offline usage for data filtering purposes after the experiment is possible by usage of the transputer host interface to a SUN workstation. The effect of electronic noise was studied. In case of a signal threshold of 3 times the R.M.S. value of the electronic noise, a fraction of 40.3% of beam background events can be rejected while 95.0% of physics events pass the classification.

References

[1] BELLE Collaboration, K. Abe et al., BELLE Technical Design Report, KEK Report 95-1, 1995.
[2] KEK B-Factory Design Report, KEK Report 95-7, 1995.
[3] The maximum BELLE DAQ readout value is limited by (a) the performance of the readout CPUs and (b) the bandwidth of the data transfer interfaces. The value of $R = 500$ Hz is valid for an interrupt driven sequence, including the whole data path from frontend readout up to storing onto a tape device. Details are given in [1], Ch. 10.1.1.
[4] R. Marshall, Z. Phys. C 26 (1984) 291.
[5] R. Itoh, P.C. Kim, QQ Quick Reference Documentation, available from http://bsunsrv1.kek.jp/ ∼ software/qq/.
[6] S. Uehara, TREPS: a Monte-Carlo event generator for two-photon processes at e^+e^- colliders, KEK Report 96-11, 1996.
[7] S.K. Sahu, A generator for study of background due to beam-gas interaction at KEK-B factory, KEK Report 97-10, 1997.
[8] GEANT 3.21 Detector description and simulation tool, CERN Program Library Long Writeup W5013, 1994.
[9] R. Itoh, BASF User's Manual, An introduction to the BELLE data analysis system, BELLE Note #161, 1996.
[10] CLEO Collaboration, R. Balest et al. A Study of Jet Production Rates in the Four Flavor Continuum and a Test a QCD, CLEO CONF preprint 94-28, presented at ICHEP94 (Ref. GLS0168), 1994.

[11] P. Söding, B. Wiik, G. Wolf, The first evidence for three-jet events in e⁺e⁻ collisions at PETRA – First Direct Observation of the Gluon, DESY preprint 96–193, 1996.

[12] TASSO Collaboration, W. Braunschweig et al., Phys. Lett. 214 B (1988) 286.

[13] CLEO Collaboration, M.S. Alam et al., Shape studies of quark jets vs gluon jets at $\sqrt{s} = 10$ GeV, CLEO preprint CLNS 92/1145, 1992.

[14] CLEO Collaboration, M.S. Alam et al., Study of gluon versus quark fragmentation in $\Upsilon \to gg\gamma$ and $e^+e^- \to q\bar{q}\gamma$ Events at $\sqrt{s} = 10$ GeV, CLEO preprint CLNS 96/1452, 1996.

[15] CLEO Collaboration, M.S. Alam et al., Two-photon production of charged pion and kaon pairs, preprint CLNS 94/1274, 1994.

[16] T. Oshima, S. Suito, A. Sugiyama, S. Suzuki, N. Haba, Examination of T/CP Invariance in the $e^+e^- \to \tau^+\tau^-$ reaction, preprint Nagoya University DPNU-97-46, BELLE Note #190, 1997.

[17] S. Uehara, Estimation of beam-background tracks in the BELLE CDC using GELHAD, BELLE Note #167, 1997.

[18] M. Lades, J.C. Vorbrügen, P. Rolf, R.P. Wütz, Recognizing faces with a transputer farm, in: T.S. Durrani, W.A. Sandham, J.J. Soraghan, S.M. Forbes (Eds.), Applications of Transputers, IOS Press, Amsterdam-Oxford-Washington-Tokio, 1991, p. 148.

[19] C. Fukunaga, M. Tanaka, H. Ikeda, Y. Igarashi, Nucl. Instr. and Meth. A 342 (1994) 175.

[20] T. Korhonen, H. Sakamoto, Y. Watase, The VENUS parallel processor trigger, in: Y. Watase, F. Abe (Eds.), Proc. Computing in High Energy Physics, 1991, Universal Academy Press, Tokyo, 1991.

[21] SGS Thomson Microelectronics, The transputer databook, INMOS Databook Series, 1992.

[22] T. Kohonen, Self-Organizing Maps, 2nd Extended ed., Springer Series in Information Sciences, vol. 30, Springer, Berlin-Heidelberg-New York, 1995, 1997.

[23] J.S. Lange, P. Schönmeier, H. Freiesleben, Nucl. Instr. and Meth. A 389 (1997) 74.

[24] T. Korhonen, Performance Analysis of a Real-Time Parallel Data Acquisition and Processing System at the TRISTAN VENUS Experiment, in: S. Noguchi, M. Ishizuka, M. Ota (Eds.), Transputer/occam-2 Japan 6, IOS Press, 1994, p. 193.

[25] D. Pountain, D. May, A Tutorial Introduction to OCCAM Programming, BSP Professional Books, Oxford, 1987.

[26] LeCroy 1810 FASTBUS calibration and trigger module, Data Sheet, available from http://www.lecroy.com/lrs/dsheets/1810.htm.

[27] G.B. Dalton, R.A.C. Croft, G. Efstathiou, W.J. Sutherland, S.J. Maddox, M. Davis, The two-point correlation function of rich clusters of galaxies: results from an extended APM cluster redshift survey. preprint astro-ph/9407076, available from http://xxx.lanl.gov.

[28] J.S. Lange, P. Hermanowski, H. Freiesleben, Nucl. Instr. and Meth. A 389 (1997) 214.

[29] B. Fritzke, Neural Networks 7 (1994) 1441.

[30] T. Martinetz, K. Schulten, A "neural-gas" network learns topologies, in: T. Kohonen, K. Mäkisara, O. Simula, J. Kangas (Eds.), Artificial Neural Networks, North-Holland, Amsterdam, 1991, p. 397.

[31] J.S. Lange, Th. Degener, J. Lüdemann, M. Kunze, M. Steinke, W. Potthast, Cluster-Gravitation – An extension to the Kohonen map for the identification of pp-Bremsstrahlung at COSY. New Computing Techniques in Physics Research IV, World Scientific, Singapore, 1995, p. 653.

[32] J.S. Lange, H. Freiesleben, A Parameter-Free Non-Growing Self-Organizing Map Based upon Gravitational Principles: Algorithm and Applications, Lecture Notes in Computer Science, vol. 1112, Springer, Berlin, 1996, p. 827.

[33] BABAR Collaboration, D. Boutigny et al., BABAR Technical Design Report, Report SLAC-R-0457, 1995.

[34] D. Goldner, Online-Datenaufbereitung fuer die Hardwarerealisierung eines kuenstlichen neuronalen Netzwerkes als Level-2 Trigger im H1-Experiment, Ph. D. Thesis, University Dortmund, 1996.

[35] T.T. Tran, S. Schiek, G. Schmidt, H.J. Behrend, G. Franke, M. Kander, K.D. Kammeyer, Application of a High Speed Analog Neural Network Chip in the First Level RZ-Trigger of the H1-Experiment at HERA, New Computing Techniques in Physics Research IV, World Scientific, Singapore, 1995, p. 365.

[36] L. Lönnblad, C. Peterson, T. Rögnvaldsson, JETNET 2.0 program, Comp. Phys. Commun. 70 (1992) 167.

[37] R. Berlich, M. Kunze, Nucl. Instr. and Meth. A 389 (1997) 274.

[38] B. Denby, Nucl. Instr. and Meth. A 389 (1997) 8.

[39] For an overview of definitions see the chapter Jet Variables in the CERN Particle Detector Briefbook, http://www.cern.ch/Physics/ParticleDetector/Brief Book/.

[40] H. Ozaki, Electronic noise effects on SVD readout, BELLE Note #173, 1997.